Approximate duality*

Clint Scovel, Don Hush, and Ingo Steinwart
CCS-3
Los Alamos National Laboratory
jcs@lanl.gov
dhush@lanl.gov
ingo@lanl.gov

September 2, 2005

Abstract

In this paper we extend the Lagrangian duality theory for convex optimization problems to incorporate approximate solutions. In particular, we generalize well-known relationships between minimizers of a convex optimization problem, maximizers of its Lagrangian dual, saddle points of the Lagrangian, Kuhn-Tucker vectors, and Kuhn-Tucker conditions to incorporate approximate versions.

1 Introduction

Duality theory provides a rich framework for the development of solution methods for convex optimization problems. Key components of this theory include a primal optimization problem, a Lagrangian defined on the space of primal and dual variables, a dual optimization problem defined on the space of dual variables (i.e. Lagrange multipliers), and Kuhn–Tucker conditions defined on the space of primal and dual variables. Solution methods include *primal methods* which work directly in the space of primal variables; *dual methods* which solve a dual optimization problem and then construct a primal solution from a dual solution; and *primal–dual methods* which solve for the primal and dual variables simultaneously. Dual and primal–dual methods are designed to find points that either satisfy the Kuhn–Tucker conditions or correspond to a saddle of the Lagrangian. Although much of existing duality theory assumes that these methods produce exact solutions, in practice it is more common to produce approximate solutions. Indeed, many practical algorithms do not converge to an exact solution in a finite number of iterations and therefore produce approximate solutions. Moreover, a simpler algorithm which produces an approximate solution is often utilized instead of a more complex algorithm which, in principle, produces an exact solution. More generally, practical algorithms operate with finite precision arithmetic and therefore the accuracy of their solutions is limited. Thus there is a need for a duality theory for approximate solutions. In particular when a dual method is used there is a need to know how best to construct an approximate primal solution from an approximate dual solution, and when a primal–dual method is used there is a need to relate the accuracy with which the Kuhn–Tucker conditions or the saddle point problem are solved to the accuracy of the corresponding approximate primal solution. Although there has been some work in this direction (See e.g. [1] for an approximate

*Los Alamos National Laboratory Report LAUR-05-6755

Kuhn–Tucker Theorem and [2] for a variational principle for approximate minimizers) we consider a general treatment of duality theory for approximate solutions. Consequently, this paper is written as an approximate version of Chapter 28 in "Convex Analysis" by Rockafellar [3] on "Ordinary Convex Programs and Lagrange Multipliers". Although we extend to optimization in Hausdorff locally convex topological vector spaces we did not extend to an infinite number of constraints. We suspect that such an extension is straightforward along the lines of [4].

2 Main results

We will state and prove an approximate version of the Kuhn-Tucker theorem and other related results as presented in [3, Chapter 28]. However, first we need to define terminology. Let X be a Hausdorff locally convex topological vector space, and consider a nonempty closed convex set $C \subset X$, a set of lower semi-continuous convex functions $f_i : X \to \mathbb{R}, i = 0,..,r$, and a set of continuous affine functions $f_i : X \to \mathbb{R}, i = r+1,..,m$. Throughout we make the following assumption:

$$\text{There exists a point } x \in C \text{ where all of the functions } f_1,..,f_r \text{ are continuous.} \tag{1}$$

We define a convex programming problem (P) as follows:

$$(P): \quad \begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & x \in C \\ & f_1(x) \le 0, ..., f_r(x) \le 0 \\ & f_{r+1}(x) = 0, ..., f_m(x) = 0. \end{array} \tag{2}$$

Let us define $C_i = \{x \in X : f_i(x) \le 0\}, i = 1, ..., r$, and $C_i = \{x \in X : f_i(x) = 0\}, i = r+1, ..., m$, and $C_0 = C \cap C_1 \cap \cdots \cap C_m$. Throughout we assume that C_0 is nonempty. We define the optimal value

$$\nu \ := \ \inf_{x \in C_0} f_0(x)$$

for the convex programming problem (P). For $\epsilon \ge 0$, we define the set of ϵ-minimizers of (P) as

$$\mathcal{O}_\epsilon(P) = \{x \in C_0 : f_0(x) \le \inf_{x' \in C_0} f_0(x') + \epsilon\}. \tag{3}$$

Let $E_r = \{\lambda \in \mathbb{R}^m : \lambda_i \ge 0, i = 1, ..., r\}$ and define the shorthand

$$h_\lambda(x) \ := \ f_0(x) + \lambda_1 f_1(x) + \cdots \lambda_m f_m(x).$$

We define the set of ϵ-Kuhn-Tucker vectors for (P) by

$$KT_\epsilon \ := \ \Big\{ \lambda \in E_r : \inf_C h_\lambda \ge \nu - \epsilon > -\infty \Big\}. \tag{4}$$

Let X^* denote the topological dual to X and consider the ϵ-subdifferential of a convex function h at x defined by

$$\partial_\epsilon h(x) \ := \ \{x^* \in X^* : h(y) \ge h(x) + x^*(y - x) - \epsilon, \quad \forall y \in X\}$$

where $\partial_\epsilon h(x) := \emptyset$ when $h(x)$ is not finite. We note the following important facts concerning the approximate subdifferential of sums of functions: If h_1 and h_2 are lower semicontinuous proper convex functions it is well known that

$$\partial_\epsilon(h_1 + h_2) \supset \bigcup_{\epsilon_1, \epsilon_2 \ge 0, \ \epsilon_1 + \epsilon_2 = \epsilon} \{\partial_{\epsilon_1} h_1 + \partial_{\epsilon_2} h_2\}. \tag{5}$$

However, if in addition there exists a point $x \in X$ where both h_1 and h_2 are finite and one of them is continuous, then [5, Theorem 2.8.3] shows that

$$\partial_\epsilon(h_1 + h_2) = \bigcup_{\epsilon_1,\epsilon_2 \geq 0,\ \epsilon_1+\epsilon_2=\epsilon} \{\partial_{\epsilon_1} h_1 + \partial_{\epsilon_2} h_2\} \tag{6}$$

Let $\partial := \partial_0$ denote the usual subdifferential. Also, let dh denote the differential of a function h. For affine continuous functions h we know that we have

$$h(x) = x^*(x) + h(0)$$

for some $x^* \in X^*$ and that

$$\partial h(x) = dh(x) = x^*, \quad \forall x \in X.$$

For a subset $S \subset X$ we let $\delta_S(x)$ denote the indicator function of the set S, i.e.

$$\delta_S(x) \ := \begin{cases} 0, & x \in S, \\ \infty, & x \notin S. \end{cases}$$

We define the set of points which satisfy the ϵ-Kuhn-Tucker conditions as

$$KTC_\epsilon \ := \left\{ (x,\lambda) \in C_0 \times E_r : \begin{array}{c} -\epsilon \leq \sum_{i=1}^m \lambda_i f_i(x), \\ \exists \epsilon_C \geq 0,\ \epsilon_i \geq 0, i = 0,..,r, \text{ such that } \sum_{i=0}^r \epsilon_i + \epsilon_C \leq \epsilon \text{ and} \\ 0 \in \partial_{\epsilon_0} f_0(x) + \sum_{i=1}^r \partial_{\epsilon_i}(\lambda_i f_i)(x) + \sum_{i=r+1}^m \lambda_i df_i(x) + \partial_{\epsilon_C}\delta_C(x) \end{array} \right\} \tag{7}$$

We define the Lagrangian on $X \times \mathbb{R}^m$ to be

$$L(x,\lambda) \ := \begin{cases} h_\lambda(x) & \lambda \in E_r, x \in C, \\ -\infty & \lambda \notin E_r, x \in C, \\ \infty & x \notin C. \end{cases} \tag{8}$$

Since C is nonempty, it follows that

$$\inf_{x \in X} L(x,\lambda) = \begin{cases} \inf_C h_\lambda, & \lambda \in E_r, \\ -\infty & \lambda \notin E_r. \end{cases} \tag{9}$$

Moreover, we note the useful identity

$$\sup_{\lambda \in \mathbb{R}^m} L(x,\lambda) = f_0(x) + \delta_{C_0}(x), \tag{10}$$

from which it follows that

$$\nu = \inf_{x \in X} \sup_{\lambda \in \mathbb{R}^m} L(x,\lambda). \tag{11}$$

We define the set of ϵ-saddle points of L to be

$$Sad_\epsilon \ := \ \big\{(x,\lambda) \in X \times \mathbb{R}^m : L(x,\lambda') - \epsilon \leq \ L(x,\lambda) \ \leq L(x',\lambda) + \epsilon, \quad \forall (x',\lambda') \in X \times \mathbb{R}^m\big\}. \tag{12}$$

We now prove an approximate version of [3, Theorem 28.3] which provides the basic connections between the sets $\mathcal{O}_\epsilon(P)$, KT_ϵ, KTC_ϵ, and Sad_ϵ.

Theorem 2.1 *For all* $0 \leq \epsilon < \infty$ *we have:*

i) $Sad_\epsilon \subset \mathcal{O}_{2\epsilon}(P) \times KT_{2\epsilon}$ *and* $\mathcal{O}_\epsilon(P) \times KT_\epsilon \subset Sad_{2\epsilon}$.

ii) $Sad_\epsilon \subset KTC_{2\epsilon}$ *and* $KTC_\epsilon \subset Sad_{2\epsilon}$.

iii) $\mathcal{O}_\epsilon(P) \times KT_\epsilon \subset KTC_{2\epsilon}$ *and* $KTC_\epsilon \subset \mathcal{O}_{2\epsilon}(P) \times KT_{2\epsilon}$.

Remark 2.2 We note that Theorem 2.1 establishes a connection between approximate primal-dual solutions and approximate primal solutions. That is, if the primal-dual method produces a $(x, \lambda) \in Sad_\epsilon$ or $(x, \lambda) \in KTC_\epsilon$, then $x \in \mathcal{O}_{2\epsilon}(P)$.

Proof: We proceed by developing an intermediate set I_ϵ and proving the intermediate assertions $I_\epsilon \subset Sad_\epsilon \subset I_{2\epsilon}$, $I_\epsilon \subset \mathcal{O}_\epsilon(P) \times KT_\epsilon \subset I_{2\epsilon}$, and $I_\epsilon \subset KTC_\epsilon \subset I_{2\epsilon}$. The theorem then follows directly. Define the intermediate set

$$I_\epsilon \ := \ \left\{(x, \lambda) \in C_0 \times E_r : f_0(x) \leq \inf_C h_\lambda + \epsilon\right\} \tag{13}$$

and suppose that $(x, \lambda) \in I_\epsilon$. Then $x \in C_0$ and the identity (10) implies that $L(x, \lambda') \leq f_0(x)$ for all $\lambda' \in \mathbb{R}^m$. Consequently, the definition (13) of I_ϵ yields

$$L(x, \lambda') - \epsilon \leq f_0(x) - \epsilon \leq \inf_C h_\lambda.$$

In addition, since $\lambda \in E_r$, the identity (9) implies that $\inf_C h_\lambda \leq L(x, \lambda)$ so that we obtain

$$L(x, \lambda') - \epsilon \leq L(x, \lambda).$$

On the other hand, since $\lambda \in E_r$, the identity (9) also implies that

$$\inf_C h_\lambda \leq L(x', \lambda), \quad \forall x' \in X$$

and since $x \in C_0$ the definition (13) of I_ϵ yields

$$L(x, \lambda) \leq f_0(x) \leq \inf_C h_\lambda + \epsilon \leq L(x', \lambda) + \epsilon.$$

Therefore $I_\epsilon \subset Sad_\epsilon$. Moreover, $\lambda \in E_r$ implies that for all $x' \in C_0$ we have $h_\lambda(x') \leq f_0(x')$ so that

$$\inf_C h_\lambda \leq \inf_{C_0} h_\lambda \leq \inf_{C_0} f_0 = \nu \leq f_0(x) \leq \inf_C h_\lambda + \epsilon$$

from which we conclude that

$$-\infty < f_0(x) \leq \nu + \epsilon$$

and therefore

$$\inf_C h_\lambda \geq \nu - \epsilon > -\infty.$$

Consequently, $I_\epsilon \subset \mathcal{O}_\epsilon(P) \times KT_\epsilon$. In addition,

$$f_0(x) \leq \inf_C h_\lambda + \epsilon \leq h_\lambda(x) + \epsilon = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \epsilon \leq f_0(x) + \epsilon$$

so that

$$-\epsilon \leq \sum_{i=1}^m \lambda_i f_i(x) \quad \text{and} \quad h_\lambda(x) \leq \inf_C h_\lambda + \epsilon.$$

The latter inequality is equivalent to

$$0 \in \partial_\epsilon (h_\lambda + \delta_C)(x). \tag{14}$$

Using the assumption (1) and the sum formula (6) we therefore obtain the existence of $\epsilon_i \geq 0, i = 0, m$, $\epsilon_C \geq 0$, $\sum_{i=0}^m \epsilon_i + \epsilon_C = \epsilon$ such that

$$0 \in \partial_{\epsilon_0} f_0(x) + \partial_{\epsilon_1}(\lambda_1 f_1)(x) + \cdots \partial_{\epsilon_m}(\lambda_m f_m)(x) + \partial_{\epsilon_C} \delta_C(x).$$

For the affine functions $f_i, i = r+1, m$, $\partial_{\epsilon_i}(\lambda_i f_i) = \partial(\lambda_i f_i) = \lambda_i df_i$, so that we conclude that $I_\epsilon \subset KTC_\epsilon$.

Now suppose that $(x, \lambda) \in Sad_\epsilon$. The inequality (12) implies that

$$\sup_{\lambda' \in \mathbb{R}^m} L(x, \lambda') - \epsilon \leq L(x, \lambda) \leq \epsilon + \inf_{x' \in X} L(x', \lambda)$$

and therefore the identities (10) and (9), and the fact that $f_0(x) \in \mathbb{R}$ imply that

$$-\infty < f_0(x) + \delta_{C_0}(x) - \epsilon \leq \begin{cases} \inf_C h_\lambda, & \lambda \in E_r, \\ -\infty & \lambda \notin E_r \end{cases} < \infty \quad .$$

Therefore we conclude that $x \in C_0$, $\lambda \in E_r$, and

$$f_0(x) \leq \inf_C h_\lambda + 2\epsilon.$$

That is, $Sad_\epsilon \subset I_{2\epsilon}$ and we have established $I_\epsilon \subset Sad_\epsilon \subset I_{2\epsilon}$.

Now suppose that $(x, \lambda) \in \mathcal{O}_\epsilon(P) \times KT_\epsilon$. Then $x \in C_0$, $\lambda \in E_r$, and

$$f_0(x) \leq \nu + \epsilon \text{ and } \inf_C h_\lambda \geq \nu - \epsilon$$

and therefore

$$f_0(x) \leq \inf_C h_\lambda + 2\epsilon$$

and so we conclude that $\mathcal{O}_\epsilon(P) \times KT_\epsilon \subset I_{2\epsilon}$ thus establishing $I_\epsilon \subset \mathcal{O}_\epsilon(P) \times KT_\epsilon \subset I_{2\epsilon}$.

Now suppose that $(x, \lambda) \in KTC_\epsilon$. It is well known that since C is closed and convex δ_C is a lower semicontinuous proper convex function. Moreover, the relation (5) applied to the subdifferential relation of KTC_ϵ implies that $0 \in \partial_\epsilon(h_\lambda + \delta_C)(x)$ which in turn implies that $h_\lambda(x) \leq \inf_C h_\lambda + \epsilon$. Moreover $-\epsilon \leq \sum_{i=1}^m \lambda_i f_i(x)$ implies that

$$h_\lambda(x) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) \geq f_0(x) - \epsilon$$

so that we obtain

$$f_0(x) \leq \inf_C h_\lambda + 2\epsilon$$

and conclude $KTC_\epsilon \subset I_{2\epsilon}$ thus establishing $I_\epsilon \subset KTC_\epsilon \subset I_{2\epsilon}$. ∎

We obtain as a corollary the following approximate version of the Kuhn-Tucker Theorem. We note that Strodiot et al. [1] prove a stronger version of ii) in $\mathbb{R}^n$.

Corollary 2.3 *For all $0 \leq \epsilon < \infty$ we have:*

i) If KT_ϵ is not empty, then for all $x \in \mathcal{O}_\epsilon(P)$ there exists a λ such that $(x, \lambda) \in Sad_{2\epsilon}$. Conversely, for a fixed x, if there exists a λ such that $(x, \lambda) \in Sad_\epsilon$ then $x \in \mathcal{O}_{2\epsilon}(P)$.

ii) If KT_ϵ is not empty, then for all $x \in \mathcal{O}_\epsilon(P)$ there exists a λ such that $(x, \lambda) \in KTC_{2\epsilon}$. Conversely, for a fixed x, if there exists a λ such that $(x, \lambda) \in KTC_\epsilon$ then $x \in \mathcal{O}_{2\epsilon}(P)$.

We now prove an approximate version of [3, Theorem 28.4] which shows how the optimal value ν relates to the value of the Lagrangian at approximate minimizers and aproximate Kuhn-Tucker vectors.

Theorem 2.4 *For all $0 \le \epsilon < \infty$ we have:*

i) $(x, \lambda) \in \mathcal{O}_\epsilon(P) \times KT_\epsilon$ implies that

$$\nu - \epsilon \le L(x, \lambda) \le \nu + \epsilon.$$

ii) $\lambda \in KT_\epsilon$ if and only if $\inf_{x \in X} L(x, \lambda) \ge \nu - \epsilon > -\infty$ and in this case

$$\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \le \epsilon.$$

Proof: Suppose that $(x, \lambda) \in \mathcal{O}_\epsilon(P) \times KT_\epsilon$. Then

$$L(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) \le f_0(x) \le \nu + \epsilon$$

and

$$L(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) \ge \nu - \epsilon$$

completes the proof of the assertion i). Now suppose that $\lambda \in KT_\epsilon$. The identity (9) implies

$$\inf_{x \in X} L(x, \lambda) = \inf_C h_\lambda \ge \nu - \epsilon > -\infty$$

and the indentity (11) then implies that

$$\sup_{\mathbb{R}^m} \inf_X L \ge \nu - \epsilon = \inf_X \sup_{\mathbb{R}^m} L - \epsilon$$

proving the forward part of assertion ii).

Conversely,

$$-\infty < \nu - \epsilon \le \inf_{x \in X} L(x, \lambda)$$

and the identity (9) implies that $\lambda \in E_r$ so that

$$\inf_C h_\lambda = \inf_{x \in X} L(x, \lambda) \ge \nu - \epsilon > -\infty.$$

Consequently $\lambda \in KT_\epsilon$ and assertion ii) is proved. ∎

Let us now consider the Lagrange dual problem

$$(D): \qquad \begin{array}{ll} \max & g(\lambda) \\ \text{s.t.} & \lambda \in \mathbb{R}^m. \end{array} \tag{15}$$

with criterion function

$$g(\lambda) \;:=\; \inf_{x \in X} L(x,\lambda). \tag{16}$$

The dual optimal value is defined as

$$\nu^* \;:=\; \sup_{\lambda \in \mathbb{R}^m} g(\lambda)$$

and the approximate maximizers for the dual problem are defined by

$$\mathcal{O}_\epsilon(D) \;:=\; \left\{ \lambda \in \mathbb{R}^m : \; g(\lambda) \geq \sup_{\lambda' \in \mathbb{R}^m} g(\lambda') - \epsilon \; \right\} \quad . \tag{17}$$

We note that the minmax inequality implies that

$$\nu^* \leq \nu.$$

We can now state the following important corollary to Theorem 2.4.

Corollary 2.5 *Consider the Lagrangian dual maximization problem (15) with concave criterion function defined in (16). Then $\mathcal{O}_\epsilon(D) \neq \emptyset$ for all $0 < \epsilon < \infty$ and for all $0 \leq \epsilon < \infty$ we have:*

i) $KT_\epsilon \subset \mathcal{O}_\epsilon(D)$.

ii) $KT_\epsilon \neq \emptyset$ *implies that* $\nu > -\infty$ *and* $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \leq \epsilon$.

iii) $\nu > -\infty$ *and* $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \leq \epsilon_1$ *for some* $0 \leq \epsilon_1 < \infty$ *implies that* $\mathcal{O}_\epsilon(D) \subset KT_{\epsilon+\epsilon_1}$.

Remark 2.6 We note that the assumptions $\nu > -\infty$ and $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \leq \epsilon$ are equivalent to the assumptions $\nu^* > -\infty$ and $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \leq \epsilon$.

Remark 2.7 The duality gap $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L$ can often be proven to be zero (See e.g. [6, Section 2.3.3]).

Proof of Corollary 2.5: It follows from the min-max inequality that $\sup_{\mathbb{R}^m} g = \sup_{\mathbb{R}^m} \inf_X L \leq \inf_X \sup_{\mathbb{R}^m} L = \nu$ and since C_0 is nonempty it follows that the righthand side is less than ∞. Consequently $\mathcal{O}_\epsilon(D) \neq \emptyset$ for all $0 < \epsilon < \infty$. Now let $\lambda \in KT_\epsilon$ for $0 \leq \epsilon < \infty$. Then Theorem 2.4 and the min-max inequality imply that

$$g(\lambda) = \inf_{x \in X} L(x,\lambda) \geq \nu - \epsilon = \inf_X \sup_{\mathbb{R}^m} L - \epsilon \geq \sup_{\mathbb{R}^m} \inf_X L - \epsilon = \sup_{\mathbb{R}^m} g - \epsilon$$

proving assertion i). Assertion ii) follows directly from assertion ii) of Theorem 2.4. For assertion iii), consider $\lambda \in \mathcal{O}_\epsilon(D)$ for $0 \leq \epsilon < \infty$. The assumptions $\nu > -\infty$ and $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \leq \epsilon_1$ implies that $\sup_{\mathbb{R}^m} \inf_X L > -\infty$. Therefore since $\lambda \in \mathcal{O}_\epsilon(D)$ we obtain

$$\inf_{x \in X} L(x,\lambda) = g(\lambda) \geq \sup_{\mathbb{R}^m} \inf_X L - \epsilon > -\infty.$$

Consequently we conclude from the indentity (9) that $\lambda \in E_r$ and

$$\inf_C h_\lambda = g(\lambda) \geq \sup_{\mathbb{R}^m} \inf_X L - \epsilon \geq \inf_X \sup_{\mathbb{R}^m} L - \epsilon_1 - \epsilon = \nu - \epsilon_1 - \epsilon$$

and the proof is finished. ∎

The following corollary is important to produce approximate primal solutions from approximate dual solutions.

Corollary 2.8 *Suppose that $0 \le \epsilon < \infty$. Then the following hold:*

i) Suppose that we have $\lambda \in \mathcal{O}_\epsilon(D)$. Also suppose that $\nu > -\infty$ and that $\inf_X \sup_{\mathbb{R}^m} L - \sup_{\mathbb{R}^m} \inf_X L \le \epsilon_1$ for some $0 \le \epsilon_1 < \infty$. When $\epsilon = 0$ suppose further that $\mathcal{O}_0(P)$ is nonempty. Then the set of x for which $(x, \lambda) \in KTC_\tau$ is nonempty for all $\tau \ge 2\epsilon + 2\epsilon_1$.

ii) Given a fixed $\lambda \in \mathbb{R}^m$, for every x such that $(x, \lambda) \in KTC_\epsilon$ we have $x \in \mathcal{O}_{2\epsilon}(P)$.

Proof: For the first claim, Corollary 2.5 implies that $\lambda \in KT_{\epsilon+\epsilon_1}$. It follows from $\nu > -\infty$ when $\epsilon > 0$ and from the monotonicity of $\mathcal{O}_\tau(P)$ in τ when $\epsilon = 0$ that $\mathcal{O}_{\epsilon+\epsilon_1}(P)$ is nonempty. From Theorem 2.1 iii) we can conclude that the set of x such that $(x, \lambda) \in KTC_{2\epsilon+2\epsilon_1}$ is not empty. The monotonicity of KTC_τ in τ proves the first claim. For the second, observe that for any $(x, \lambda) \in KTC_\epsilon$ it follows also from Theorem 2.1 iii) that $x \in \mathcal{O}_{2\epsilon}(P)$. ■

Remark 2.9 We note that Corollary 2.8 provides a mechanism for generating approximate solutions to the primal problem (P) from approximate solutions to its dual (D); Suppose the duality gap is zero and $\nu > -\infty$. Then given an ϵ-maximizer λ of the dual problem, Corollary 2.8(i) states that there exists an x such that (x, λ) satisfy the approximate Kuhn Tucker equations $KTC_{2\epsilon}$. Solve these equations for some x and then Corollary 2.8(ii) shows that x is a 4ϵ-minimizer of the primal problem (P).

Remark 2.10 In the above remark, we mentioned a technique for producing approximate solutions to the primal problem from approximate solutions to its dual by solving the approximate Kuhn Tucker equations *exactly*. However, this formalism allows approximate solutions to the aproximate Kuhn Tucker equations in the following way: Choose solution methods such that approximate solutions to KTC_ϵ are exact solutions to $KTC_{\epsilon'}$ for some $\epsilon' \ge \epsilon$.

References

[1] J.-J. Strodiot, V. Hien Nguyen, and N.Heukemes. ϵ-optimal solutions in nondifferentiable convex programming and some related questions. *Mathematical Programming*, 25:307–328, 1983.

[2] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[3] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[4] P.P. Varaiya. Nonlinear programming in Banach space. *SIAM Journal on Applied Mathematics*, 15:284–293, 1967.

[5] C. Zălinescu. *Convex analysis in general vector spaces.* World Scientific, New Jersey, 2002.

[6] V. Barbu and Th. Precupanu. *Convexity and Optimization in Banach Spaces.* D. Reidel Publishing Company, Dordrecht, 1986.